1. (Currently amended) A collection of software tools embodied on a computer readable medium for acquiring data from diverse sources and/or structuring the data and/or determining similarity of content for the purpose of product information management, said collection comprising:

two or more tools selected from the group consisting of a web agent creator for creating a web agent to seek out and acquire product information on the world wide web, a web agent created by the web agent creator, the web agent capable of acquiring product information from the world wide web, a web agent manager, an ontology-directed classifier capable of classifying product information, an ontology-directed extractor capable of extracting product information from content contained in unstructured textual product descriptions, and an ontology-directed matcher capable of matching product information extracted by the extractor through matching product categories and attributes.

2. (original) The collection according to claim 1, wherein:
one or more of the tools are example driven through a graphical user interface.

3. (original) The collection according to claim 1, wherein:
said web agent creator has a web browser interface and a web agent is created by navigating to a web page of interest and selecting the kind of information to be extracted from the web page.

4. (original) The collection according to claim 1, wherein:

said web agent creator includes

      a web browser user interface,

      a pattern expression discovery algorithm coupled to said user interface,

      a results editor coupled to said user interface and said pattern expression

discovery algorithm,

      an agent generator coupled to said user interface and said results editor, and

      a form value editor coupled to said user interface and said agent generator.

5. (original) The collection of claim 4, wherein:

said user interface indicates text selected by the user interface to said pattern expression

discovery algorithm, said results editor, said agent generator, and said form value editor.

6. (original) The collection of claim 4, wherein:

said pattern expression discovery algorithm is an XPath discovery algorithm,

said user interface indicates a DOM tree of text selected by the user interface to said

XPath discovery algorithm, said results editor, said agent generator , and said form value

editor.

7. (original) The collection of claim 5, wherein:

said pattern expression discovery algorithm generates a pattern expression based on the results received from the user interface and communicates that pattern expression to the results editor.

8. (original) The collection of claim 6, wherein:

said XPath discovery algorithm generates an XPath based on the DOM tree received from the user interface and communicates that XPath to the results editor.

9. (original) The collection of claim 7, wherein:

the results editor receives pattern expressions from the pattern expression discovery algorithm and accepts input from the user interface to identify the nature of the selected text.

10. (original) The collection of claim 8, wherein:

the results editor receives XPath expressions from the XPath discovery algorithm and accepts input from the user interface to identify the nature of the selected text.

11. (original) The collection of claim 8, wherein:

the form value editor receives input from the user interface and provides output to the agent generator including instructions and data to be used by the agent generated by the agent generator to fill out web based forms in order to reach the source of data to be extracted by the agent.

12. (original) The collection of claim 11, wherein:

the pattern expression discovery algorithm takes as its input a set of items corresponding

to the text highlighted by the user interface,

identifies the items, and

determines corresponding data extractor and isolator expressions.

13. (original) The collection of claim 11, wherein:

the pattern expression discovery algorithm is an XPath discovery algorithm,

the XPath discovery algorithm takes as its input a set of nodes corresponding to the text

highlighted by the user interface,

identifies locator nodes and grouping nodes based on the input set of nodes, and

determines corresponding data extractor and isolator expressions.

14. (original) The collection according to claim 12, wherein:

the corresponding data extractor and isolator expressions are used to form a navigation

map to be used by the agent to

      find all nodes that match the isolator expression, and

      for each node matching the isolator expression, find a match for each of the data

extractor expressions.

15. (original) The collection according to Claim 1, wherein:

the ontology directed classifier uses a taxonomy provided by a tree of classes and

subclasses generated using an ontology management system.

16. (original) The collection according to Claim 15, wherein:

the ontology directed classifier performs taxonomy token weighting, node weighting for

descriptions, weight propagation and normalizations, and determining the best class and

subtree of said taxonomy to which an item can be classified.

17. (original) The collection according to claim 1, wherein:

said ontology directed extractor takes unstructured text descriptions about an item as

input and produces a set of structured property values about the item as output.

18. (currently amended) A web agent creator embodied on a computer readable medium for creating a web agent to acquire product information ~~data~~ from the world wide web, said web agent creator comprising:

a web browser user interface,

a pattern expression discovery algorithm coupled to said user interface, said algorithm capable of discovering patterns of product information,

a results editor coupled to said user interface and said pattern expression discovery algorithm, said results editor capable of editing product information,

an agent generator coupled to said user interface and said results editor, said generator capable of generating said web agent having characteristics determined by said algorithm, and

a form value editor coupled to said user interface and said agent generator, said form value editor capable of setting parameters of said web agent.

19. (original) The web agent creator according to claim 18, wherein:
said user interface indicates text selected by the user interface to said pattern expression discovery algorithm, said results editor, said agent generator, and said form value editor.

20. (original) The web agent creator according to claim 18, wherein:
said pattern expression discovery algorithm is an XPath discovery algorithm,
said user interface indicates a DOM tree of text selected by the user interface to said XPath discovery algorithm, said results editor, said agent generator , and said form value editor.

21. (original) The web agent creator according to claim 19, wherein:

said pattern expression discovery algorithm generates a pattern expression based on the

results received from the user interface and communicates that pattern expression to the

results editor.

22. (original) The web agent creator according to claim 20, wherein:

said XPath discovery algorithm generates an XPath based on the DOM tree received

from the user interface and communicates that XPath to the results editor.

23. (original) The web agent creator according to claim 18, wherein:

the results editor receives pattern expressions from the pattern expression discovery

algorithm and accepts input from the user interface to identify the nature of the selected

text.

24. (original) The web agent creator according to claim 20, wherein:

the results editor receives XPath expressions from the XPath discovery algorithm and

accepts input from the user interface to identify the nature of the selected text.

25. (original) The web agent creator according to claim 18, wherein:

the form value editor receives input from the user interface and provides output to the agent generator including instructions and data to be used by the agent generated by the agent generator to fill out web based forms in order to reach the source of data to be extracted by the agent.

26. (original) The web agent creator according to claim 18, wherein:

the pattern expression discovery algorithm takes as its input a set of items corresponding to the text highlighted by the user interface,

identifies the items, and

determines corresponding data extractor and isolator expressions.

27. (original) The web agent creator according to claim 18, wherein:

the pattern expression discovery algorithm is an XPath discovery algorithm,

the XPath discovery algorithm takes as its input a set of nodes corresponding to the text highlighted by the user interface,

identifies locator nodes and grouping nodes based on the input set of nodes, and

determines corresponding data extractor and isolator expressions.

28. (original) The web agent creator according to claim 26, wherein

the corresponding data extractor and isolator expressions are used to form a navigation

map to be used by the agent to

      find all nodes that match the isolator expression, and

      for each node matching the isolator expression, find a match for each of the data

extractor expressions.

29. (currently amended) An ontology directed classifier embodied on a computer

readable medium for use with an ontology management system designed to manage

product information, said ontology directed classifier comprising:

means for receiving a product information related taxonomy as input; and

means for generating a tree of product information classes and subclasses as output for

use by the ontology management system to classify product information.

30. (original) The ontology directed classifier according to claim 29, further comprising:

means for taxonomy token weighting,

means for node weighting for descriptors

means for weight propagation and normalization, and

means for determining the best class and sub-tree of said taxonomy to which an item can

be classified.

31. (currently amended) An ontology directed extractor embodied on a computer readable medium for use with an ontology management system ~~and~~ for extracting product information attributes and their values, said ontology directed extractor, comprising:

means for receiving an unstructured text description about ~~an~~ a product ~~item~~ as input, and

means for producing a set of structured property values about the product ~~item~~ as output, wherein

said structured property values are structured by ontology relationships.

32. (cancel)

33. (currently amended) An ontology directed matcher embodied on a computer readable medium for use with an ontology management system to match similar products using product attributes and their values, said ontology directed matcher comprising:

means for describing products ~~items~~ based on a structured set of properties;

means for defining the relative importance of said properties in describing said products ~~items~~; and

means for scoring the degree of equivalence of products ~~items~~ based on said definitions.

34. (original) An ontology directed matcher according to claim 33, wherein:

said structured set of properties in defined by ontology attributes provided by the ontology management system.

35. (original) An ontology directed matcher according to claim 34, wherein:

said means for defining the relative importance of said properties is based on weight attached to a matching function for each said property that takes as input the values of said attributes defining that property for two different items and outputs a number indicating the similarity of these input values.

36. (original) An ontology directed matcher according to claim 35, wherein:

said means for scoring the degree of equivalence of items includes means for multiplying the said output values of all said matching functions by said respective weights and summing these products.

37. (original) The collection according to claim 1, further comprising:

a validation method applied to one or more tools in the collection to determine the accuracy of the tool's output by manually checking the accuracy of a statistical sampling of tool output from specific tool input.

38. (original) The collection according to claim 37, wherein:

said validation method determines an Acceptable Quality Level (AQL) as defined in standard ANSI/ASQC Z1.4-1993 by performing multiple sampling procedures at different AQLs as defined in said standard until the boundary AQL level is found below which the sampling procedure fails and above which the sampling procedure succeeds.